Evaluation of Speech Processing Schemes using Binaural Dichotic Presentation to Reduce the Effect of Masking in Hearing-Impaired Listeners

Alice N. Cheeran[1] and Prem C. Pandey[2]

[1]Biomedical Engg. Group, [2]Electrical Engg. Dept.
Indian Institute of Technology, Bombay, India
{alicenc, pcpandey}@ee.iitb.ac.in

Abstract

Binaural dichotic presentation schemes help in increasing the intelligibility of speech for persons with moderate bilateral sensorineural hearing impairment. Spectral splitting, with comb filters has helped in improving the perception of place feature. Temporal splitting, using trapezoidal fading functions, provided improvement in the perception of duration feature. Splitting of speech using time-varying comb filters, in which the bands of the comb filter are cyclically swept to have spectral and temporal splitting simultaneously, provided improvement in features of place and duration, without affecting the other features. An overall evaluation has been carried out, by conducting listening tests on seven normal hearing subjects with simulated hearing loss, on the three processing schemes: (i) spectral splitting with comb filters based on auditory critical bands designed for minimum spectral distortion and perceptual balance, (ii) temporal splitting with inter-aural switching of 20, 40, and 80 ms with trapezoidal fading functions of 70 % duty cycle and 3 ms transition, and (iii) combined spectral and temporal splitting using time-varying comb filters with sweep cycles of 20, 40, 80, 120 and 160 ms, each with 4, 8, and 16 shiftings. Test material consisted of phonetically balanced monosyllabic words in listener's native language. Hearing loss was simulated by adding broadband noise at different levels to the speech stimuli at constant short-time SNR of 10 ms. At recognition score of 60 %, spectral splitting provided an improvement of 5 dB in SNR. Combined splitting with sweep cycle of 80 ms or greater with 8 and 16 shiftings provided almost the same improvement.

1. Introduction

Sensorneural hearing impairment is characterized by frequency dependant shifts in hearing threshold, loudness recruitment, reduced frequency and temporal resolution, and increased spectral and temporal masking [1], [2]. The clarity of speech perceived reduces due to the smearing of the peaks of the speech spectrum. Increased temporal masking leads to increase in forward and backward masking of weak acoustic segments by adjacent strong ones Masking takes place primarily at the peripheral level, while integration of information takes place at higher levels in the auditory system. Earlier investigations, using schemes that split speech into two for binaural dichotic presentation, has reported to have reduced the effect of increased masking for persons with residual hearing in both the ears [3]-[6].

Lunner [3] investigated a scheme of splitting with comb filters having eight channels with constant filter bandwidth of 700 Hz, designed with complementary interpolated linear phase FIR filters. An improvement of 2 dB in speech-to-noise ratio for dichotic over diotic was reported. Chaudhari and Pandey [4] used a pair of comb filters with complementary magnitude responses based on auditory critical bandwidths [7]. The filters were designed with the consideration of relatively flat response in the pass band, high attenuation in stop bands and sharp inter-band transition. Evaluation of the scheme on bilateral sensorineural hearing-impaired subjects showed significant improvement in recognition scores and perception of consonantal features particularly the place.

Further a pair of comb filters based on 18 critical bands over 5 kHz range was designed as 256-coefficient linear phase FIR filters, with sampling rate of 10 k Sa/s, for minimum spectral distortion and perceptual balance. For design of such filters, frequency sampling technique was used in an iterative manner, treating one or two transition samples as unconstrained [8], which are adjusted to obtain the required magnitude response. Listening tests with these filters established that inter-band crossover gain adjusted within 4-6 dB of the pass band gain resulted in perceptual balance and 1 dB ripple in the pass band was found to be perceptually acceptable [9]. These filters have transition width of 78 to 117 Hz and stop band attenuation of 30 dB. Listening tests on normal subjects with simulated hearing loss and on moderate bilateral sensorineural hearing impaired, showed higher improvement in recognition scores and information transmission of consonantal features for perceptually balanced filters compared to comb filters with sharp transitions.

Jangamashetti and Pandey [5], investigated a scheme of temporal splitting, in which speech was switched between two ears using trapezoidal fading function with an inter-aural switching period of 20 ms. Evaluation on normal subjects with hearing loss simulated at different levels, for different duty cycles and transition durations, resulted in improvement of consonantal duration feature. More improvement was obtained for 70 % duty cycle with 2 and 3 ms transition durations.

In the cochlear basilar membrane the sensory cells corresponding to alternate bands are always relaxed in spectral splitting, while the sensory cells of the two ears are alternately relaxed in temporal splitting. A scheme of combined splitting was devised, in which all the sensory cells of the basilar membrane are periodically relaxed from stimulation for some time. The scheme is implemented using a pair of time varying comb filters with pre-calculated set of coefficients, which were selected in steps such that a cyclic sweeping of the pass bands occur. An experimental evaluation conducted on normal subjects with simulated hearing loss with constant sweep cycle of 20 ms provided improvement in the recognition scores and perception of place and duration features.

Presently an overall evaluation of all the three schemes of binaural dichotic presentation, spectral, temporal, and combined splitting with different processing parameters, is carried out by conducting listening tests, on normal hearing subjects with simulated hearing loss. Listening tests involved recognition of randomly presented phonetically balanced monosyllabic words.

2. Processing Schemes

Evaluation on normal subjects was done for the schemes of spectral, temporal, and combined splitting, with various processing parameters. Spectral splitting scheme (SS) with perceptually balanced comb filters based on auditory critical bands is shown in Fig. 1. The perceptually balanced comb filters with 9 pass bands were designed as 256-coefficient linear phase FIR filters. Fig 2 shows temporal splitting (TS) and the fading functions used for inter-aural switching. Based on earlier results [5], it was decided to use 70 % duty cycle and 3 ms transition duration. Constant inter-aural switching period of 20 ms was used in the earlier investigation. Presently, we have used three inter-aural switching periods, 20 ms, 40 ms and 80 ms (TS-20, TS-40, TS-80).

The schematic representation of the combined splitting scheme (CS) with time-varying comb filters is shown in Fig. 3. For an implementation with *m* shiftings, each of the time-varying comb filter contained *m* perceptually balanced comb filters, which have magnitude responses such that the pass bands of each of these comb filter pairs are shifted in a complementary manner along the frequency axis. Combined splitting is considered with sweep cycle of 20 40, 80, 120 and 160 ms, each with 4, 8, and 16 shiftings (CS-20/4, CS-20/8, CS-20/16, CS-40/4, CS-40/8, CS-40/16, CS-80/4, CS-80/8, CS-80/16, CS-120/8, CS-120/16, CS-160/8, and CS-160/16).

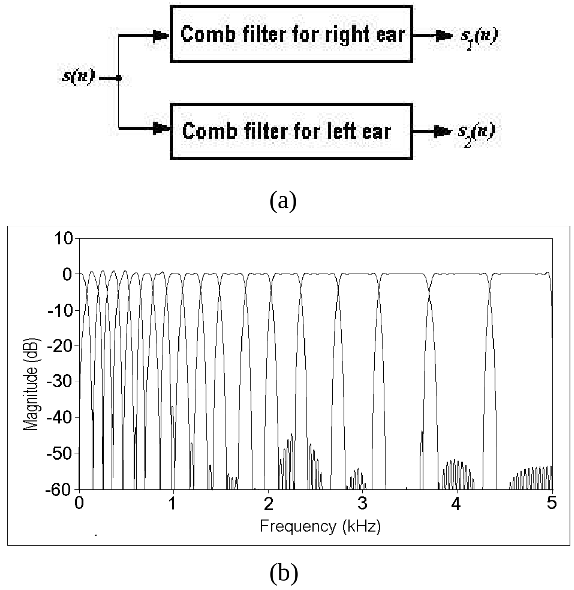


Fig 1. Spectral splitting (SC): (a) Schematic (b) Perceptually balanced comb filters

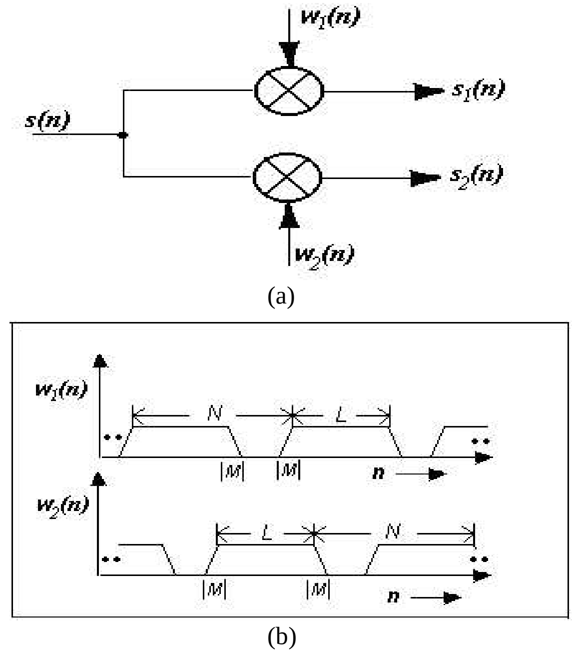
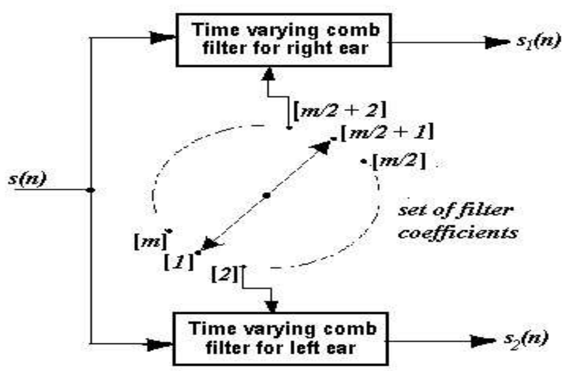
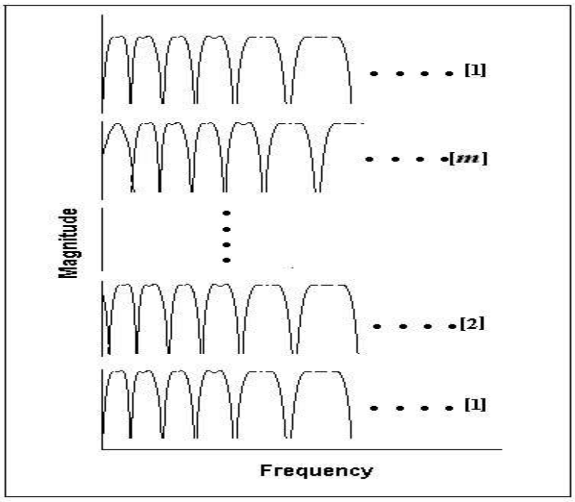


Fig. 2. Temporal splitting (TS): (a) schematic (b) trapezoidal fading functions with inter-aural switching period of *N samples*, duty cycle *L*/*N*, transition duration of *M samples.*



(a)



(b)

Fig. 3. Combined splitting (CS): (a) schematic (b) magnitude response of a time-varying comb filters

3. Experimental Evaluation

The schemes were implemented for off-line processing. Experimental evaluation was carried out through listening tests on normal subjects. Different levels of bilateral sensorineural hearing loss was simulated on normal subjects, by adding broad-band Gaussian noise with short-time signal-to-noise ratios (SNR) of 10 ms. In the present evaluation, SNRs of ∞, 3, 0, -3, -6, -9 dB were used. The schemes evaluated were binaural diotic presentation of unprocessed speech (Su) and binaural dichotic presentation of processed speech: with spectral splitting (SS), temporal splitting (TS), and combined splitting (CS) for the different processing conditions mentioned above.

The test material used for listening test consisted of words presented in a randomized order from a phonetically balanced list of 47 monosyllabic words in Marathi, the first language of the subjects who participated in the tests. This word list is in use at AYJ National Institute for Hearing Handicapped, Mumbai, for evaluating speech discrimination by the hearing impaired. The words in the list were recorded from a male speaker at 10 k samples/s and bit quantization of 16 bits, using the line-in of the PC sound card. All the words had approximately the same intensity. The recorded signals were processed off-line for the three processing schemes, with the different parameters. In the listening test set-up, a PC with sound card was used for presentation of the signals through the two output channels of the sound card, two audio amplifiers, and a pair of audiometric headphones, to the subject seated in an acoustically isolated room. After each presentation, the subject responded verbally through a microphone, and the response was entered on the PC keyboard as right or wrong by the experimenter sitting outside. In a test, each word was presented 3 times. The responses as well as the response times were tabulated. Before each test, the subject listened to the words in an order listed, to become familiar, as many times as required. Seven normal hearing subjects all male (aged 18-30) participated in the listening tests.

4. Test Results

Fig 4. shows the average recognition scores of unprocessed and processed speech for all processing conditions tested on seven normal hearing subjects for the different SNRs considered. The average recognition score for unprocessed speech reduced from 99.8 % at no noise condition to 23.9 % at -9 dB SNR. At all SNR conditions except at 3 dB, spectral splitting showed highest percentage relative improvement. For 60% recognition score spectral splitting provided an improvement of approximately 5 dB in SNR. Combined splitting with sweep cycle 80 ms with 8 and 16 shiftings resulted in recognition scores close to spectral splitting. The SNR improvement for 60% recognition score was slightly less than 5 dB SNR for combined splitting with sweep cycle of 80 ms for 8 and 16 shiftings. Combined splitting with time varying comb filters are best at 80 ms sweep cycle. The results show very less variation in recognition scores for processing with 8 and 16 shiftings. In case of temporal splitting, out of the three inter-aural switching periods (20, 40, and 80 ms), results for 20 ms were better under all SNR conditions. Further the temporal splitting appears to be effective in improving the speech perception at higher SNRs, but not at lower SNRs.

5. Conclusion

Speech processing schemes for improving the perception by reducing the spectral and temporal masking were investigated on normal persons by simulating sensorineural loss. Maximum number of subjects showed improvements with spectral splitting with perceptually balanced comb filters. Between the two intra-aural switching intervals for temporal splitting, improvements were higher for 20 ms. For combined splitting, the sweep cycle of 20 ms gave the poorest results, and best results were obtained for 80 ms.

In summary, the dichotic processing schemes implemented in binaural hearing aids should improve speech perception for persons under adverse listening conditions. Further, listening tests are conducted on moderate bilateral sensorineural hearing-impaired subjects. Combination of these schemes with multiband compression needs to be investigated.

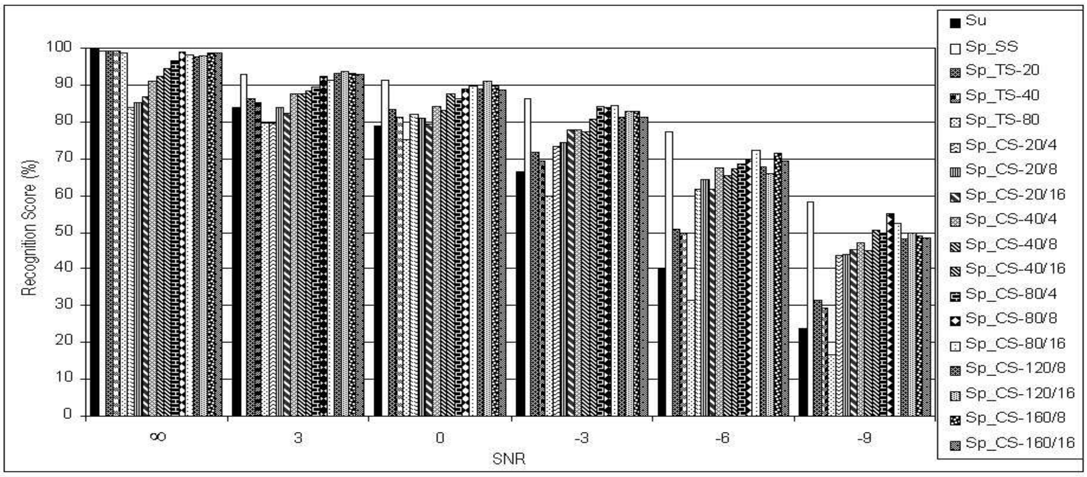


Fig 4. Percentage recognition scores for Unprocessed speech (Su) and processed speech for listening tests conducted on normal subjects for simulated hearing loss of ∞, 3, 0, -3, -6, -9 dB -SNRs, averaged over seven

6. References

[1] Moore, B. C. J., "Psychoacoustics of cochlear hearing impairment and design of hearing aids," *Proc. 16th Int. Congress Acoust. (ICA)*, Seattle, Washington, USA, pp. 2105-2108, 1998.

[2] Baer, T., and Moore, B. C. J., "Effects of spectral smearing on the intelligibility of sentences in noise," *J. Acoust. Soc. Am.*, Vol. 94, pp. 1229-1241, 1993.

[3] Lunner, T., Arlinger, S., and Hellgren, J., "8-channel digital filter bank for hearing aid use: preliminary results in monaural, diotic and dichotic modes," *Scand. Audiol.*, Suppl. 38, pp. 75-81, 1993.

[4] Chaudhari, D. S., and Pandey, P. C., "Dichotic presentation of speech signal with critical band filtering for improving speech perception," *Proc. ICASSP'98*, Seattle, Washington, USA, paper AE 3.1, 1998.

[5] Jangammashetti, D. S., and Pandey, P. C., "Inter-aural switching with different fading functions for dichotic presentation to reduce the effect of temporal masking in sensorineural hearing loss," *Proc. SCI'2000*, Orlando, Florida, USA, pp.434-439, 2000.

[6] Zwicker, E., "Subdivision of the audible range into critical bands (Frequenzgruppen)," *J. Acoust. Soc. Am.*, Vol. 33, pp. 248, 1961.

[7] Rabiner, L. R., Gold, B., and McGonegal, C. A., "An approach to the approximation problem for nonrecursive digital filters," *IEEE Trans. Audio and Electroacoustics*, Vol. AU-16, pp. 83-105, 1970.

[8] Cheeran, A. N., and Pandey, P. C., "Dichotic presentation for binaural hearing aids using perceptually balanced critical bandwidth based comb filters," *Proc. SCI'2003*, Orlando, Florida, USA, 2003